Standard Deviation vs Standard Error

Standard Deviation

Standard Deviation:

- A measure of **how spread out** the data is from the mean
- Amount of **variation** in data
- Does **NOT** tend to change as we increase sample size

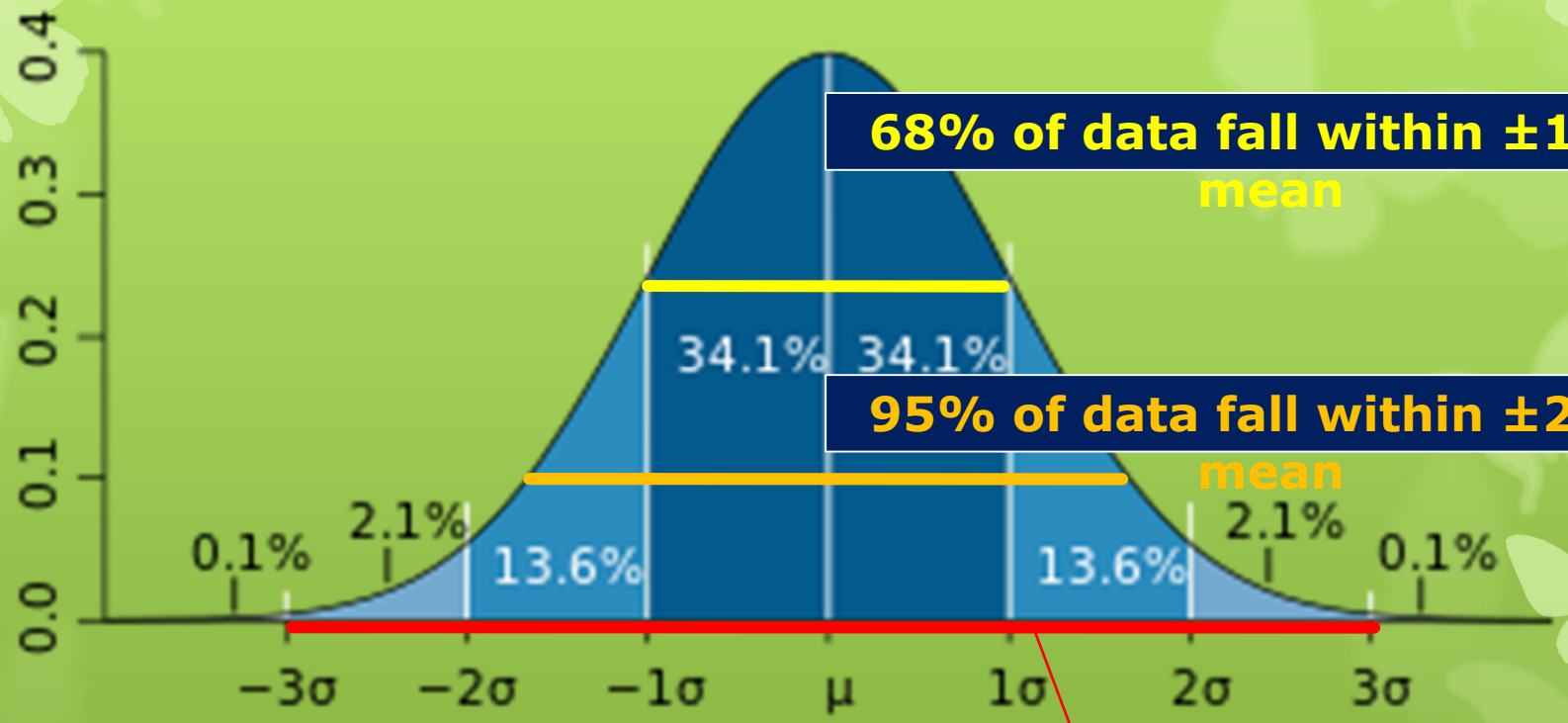
Lower standard deviation:

- Data is **closer to the mean**
  - Not as much variation around mean
- Greater likelihood that the independent variable is causing the changes in the dependent variable

Higher standard deviation:

- Data is more **spread out from the mean**
  - **Lots of variation** around mean
- More likely factors, other than the independent variable, are influencing the dependent variable

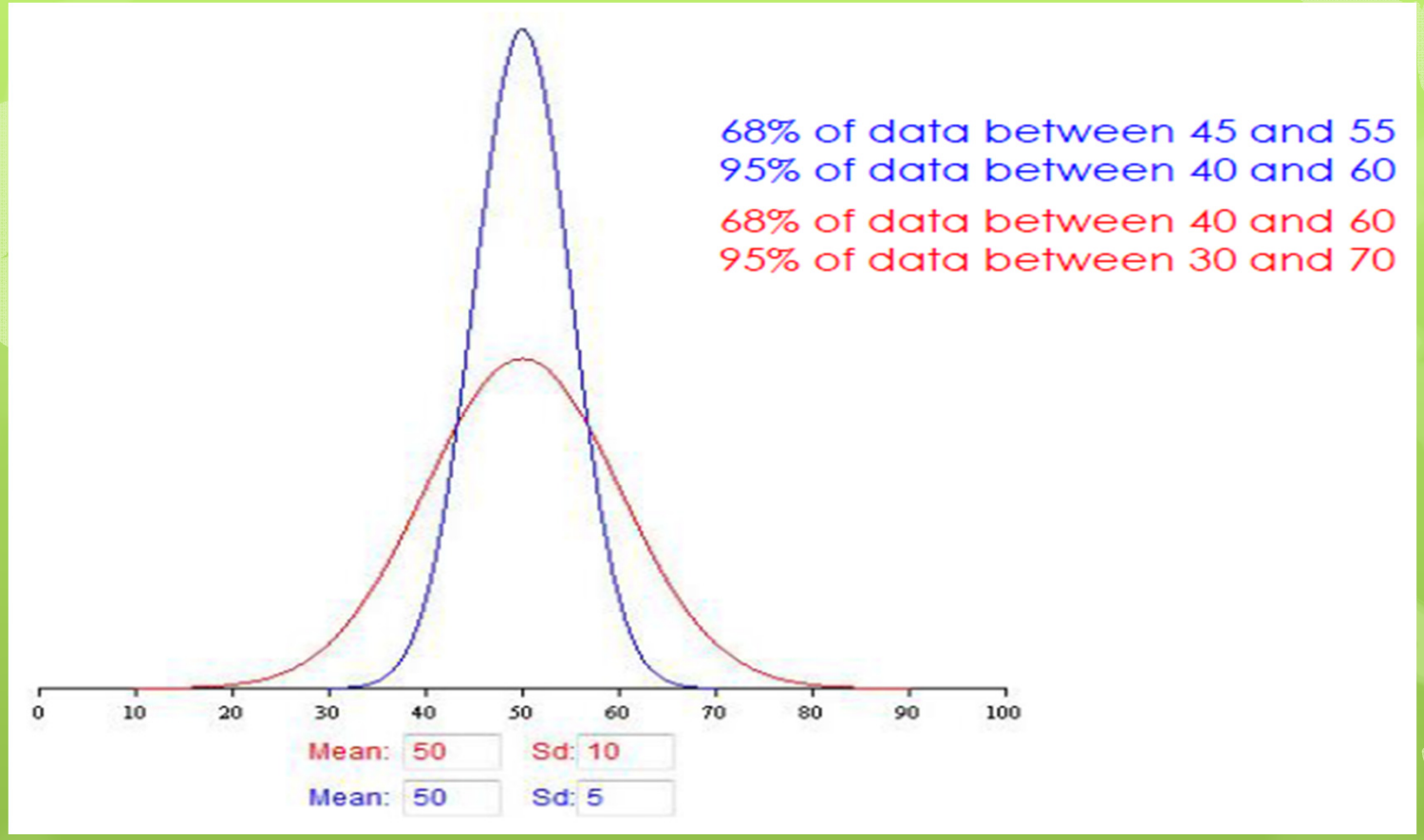
σ = standard deviation



68% of data fall within ±1s of mean

95% of data fall within ±2s of mean

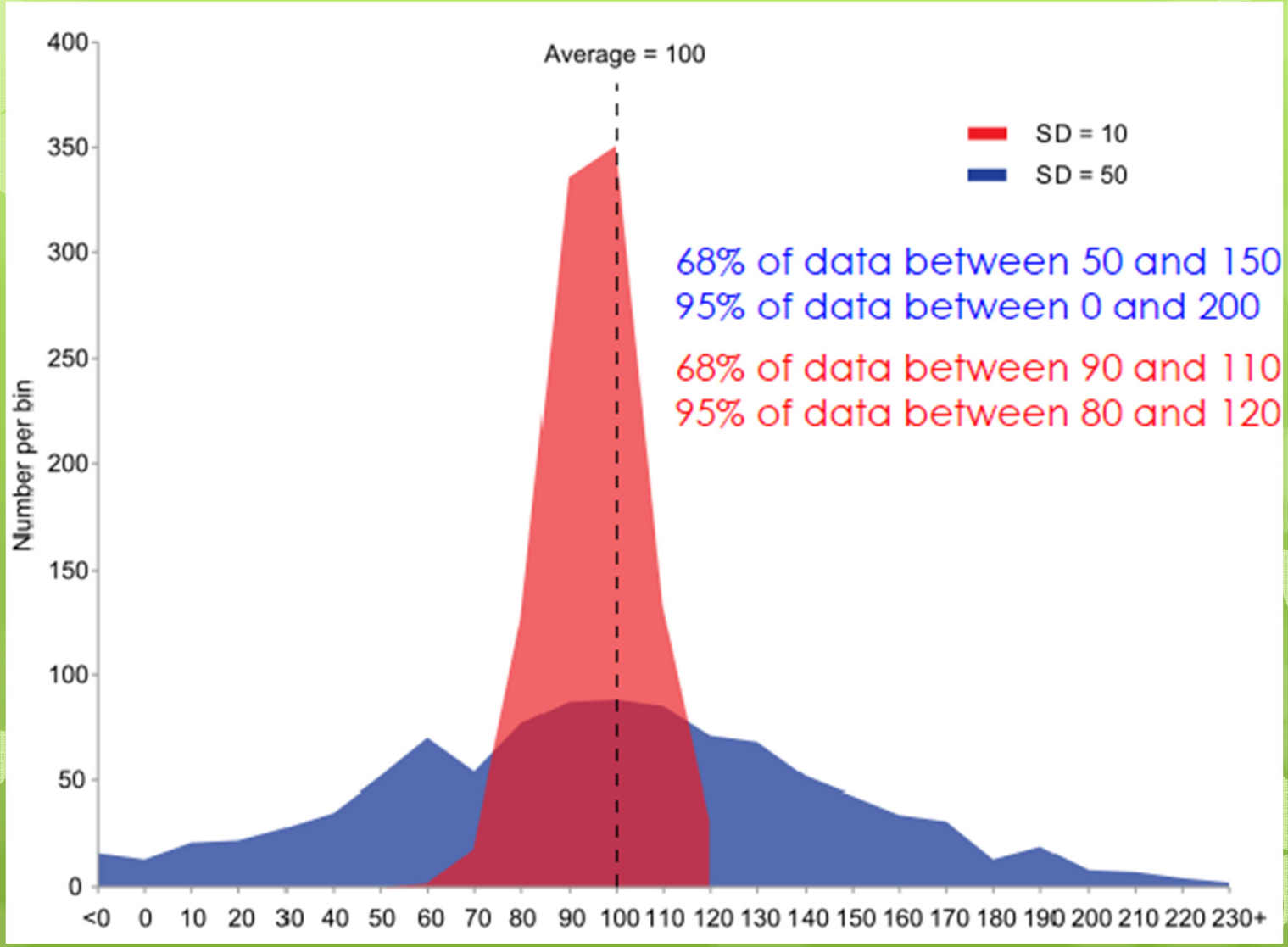
99% of data fall within ±3s of mean



- The magnitude of the standard deviation depends on the spread of the data set

- Two data sets:
  - Same mean;
  - Different standard deviation

Actual data sets aren't always so pretty…



Calculating standard deviation, s

1. Calculate the mean ($\bar{x}$)
2. Determine the difference between each data point, and the mean
3. Square the differences
4. Sum the squares
5. Divide by sample size (n) minus 1
6. Take the square root

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

Standard Error

Standard Error:

- Indication of **how well the mean of a sample ($\bar{x}$) estimates the true mean of a population (μ)**
- How **confident** we are in our estimate of the mean
  - **Large SE** = **not very** confident
  - **Small SE** = **more** confident
- Standard error **decreases** as sample size **increases**

- Measure of accuracy, if the true mean is known
  - **Accuracy** – How close a measured value is to the **actual (true) value**
- Measure of precision, if true mean is not known
  - **Precision** – How close the measured values are **to each other**.

Calculating Standard Error, SE

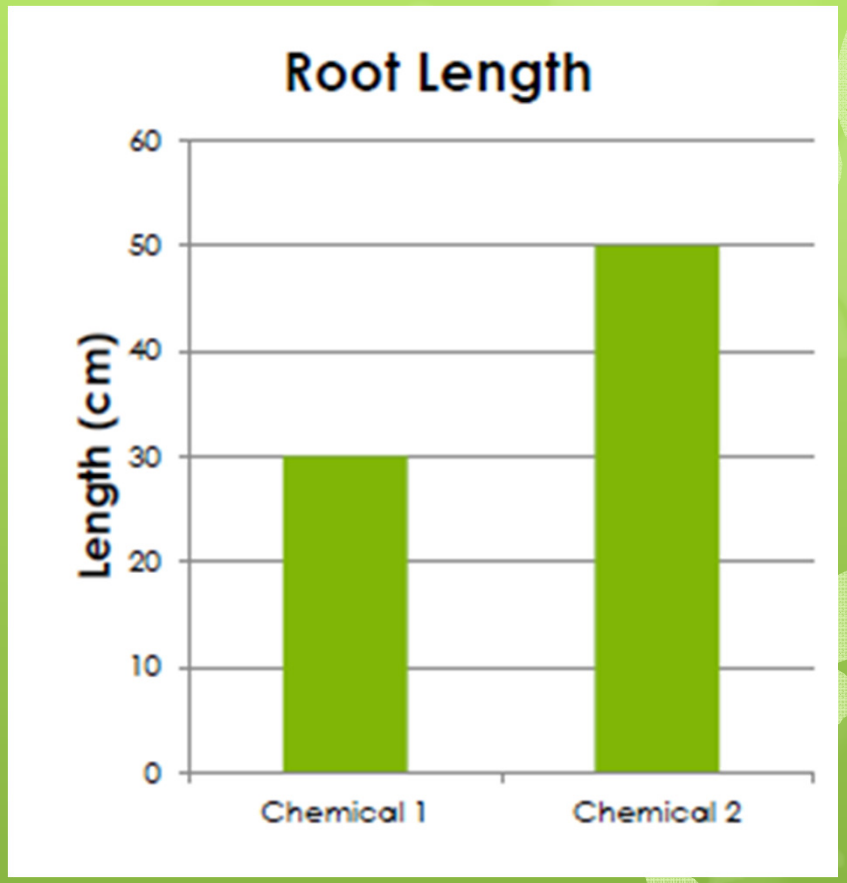
$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

1. Calculate standard deviation
2. Divide standard deviation by square root of sample size

How do we use Standard Error?

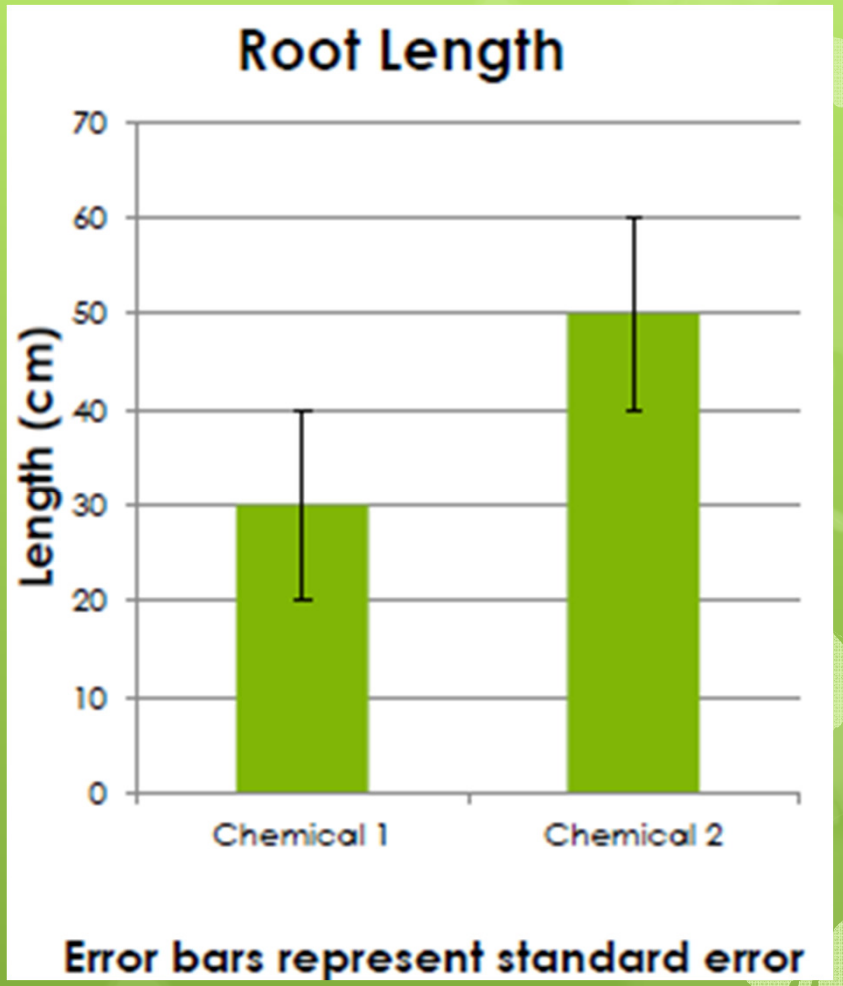
Create bar graph

- mean on Y-axis
- sample(s) on the X-axis
- chemical 1 mean = 30 cm
- chemical 2 mean = 50 cm



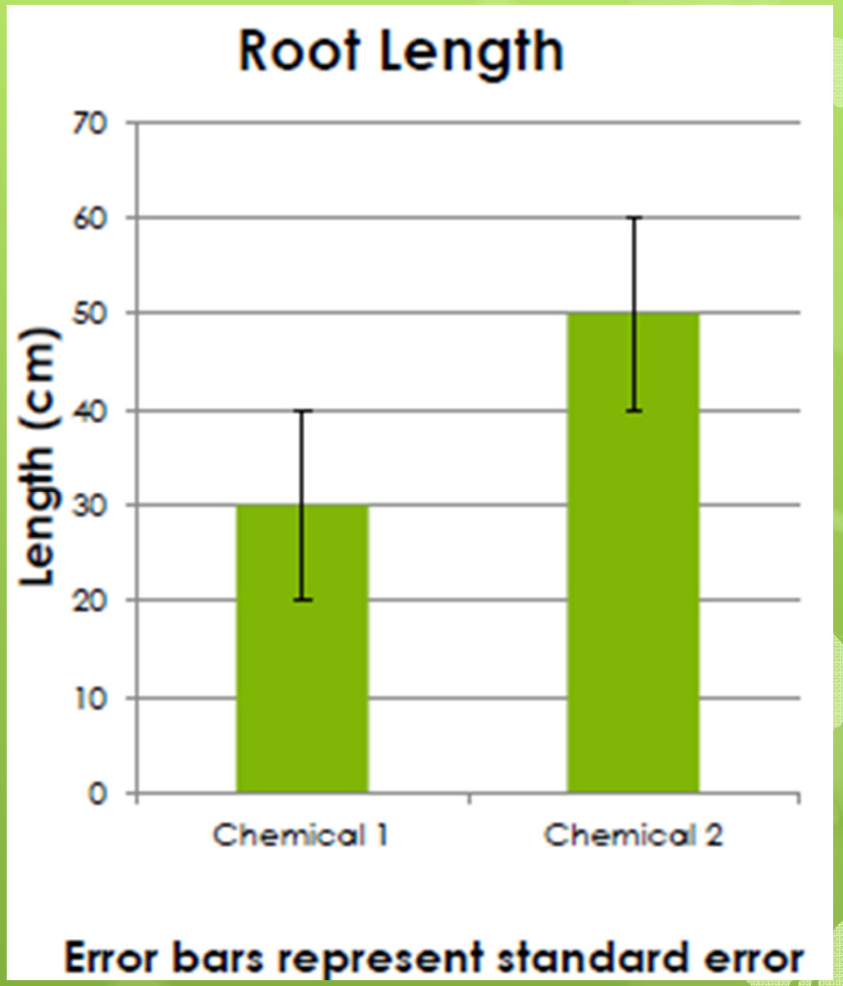
Add error bars!

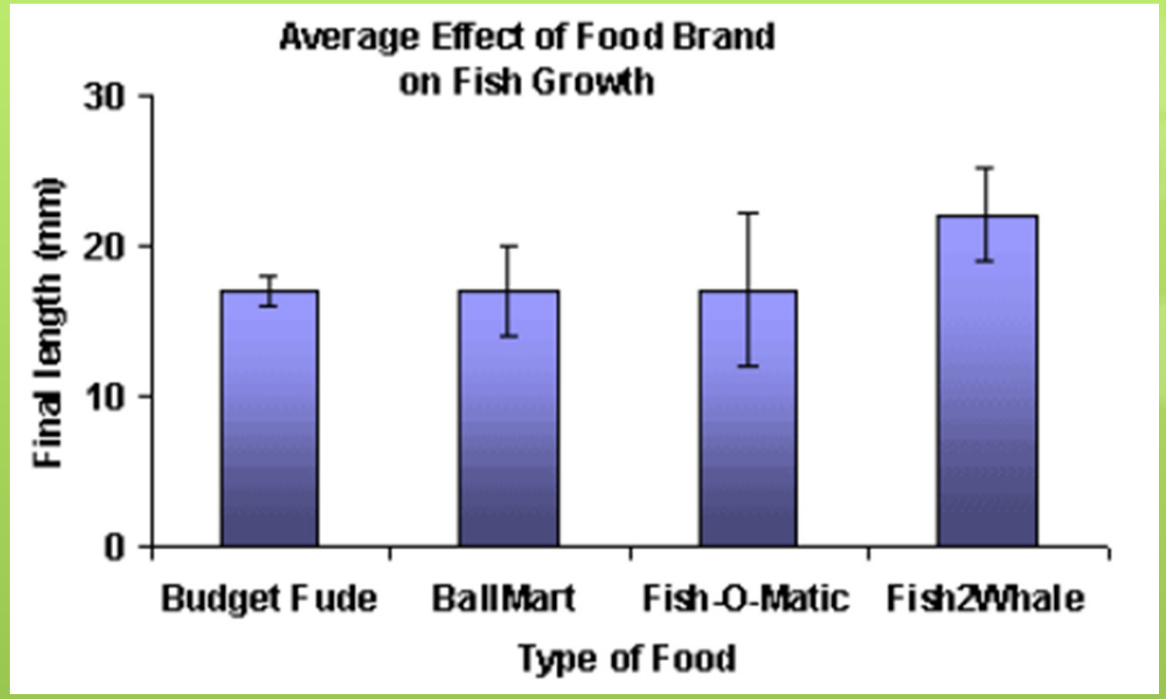
- ± SE
- Indicate in figure caption that error bars represent standard error (SE)



Analyze!

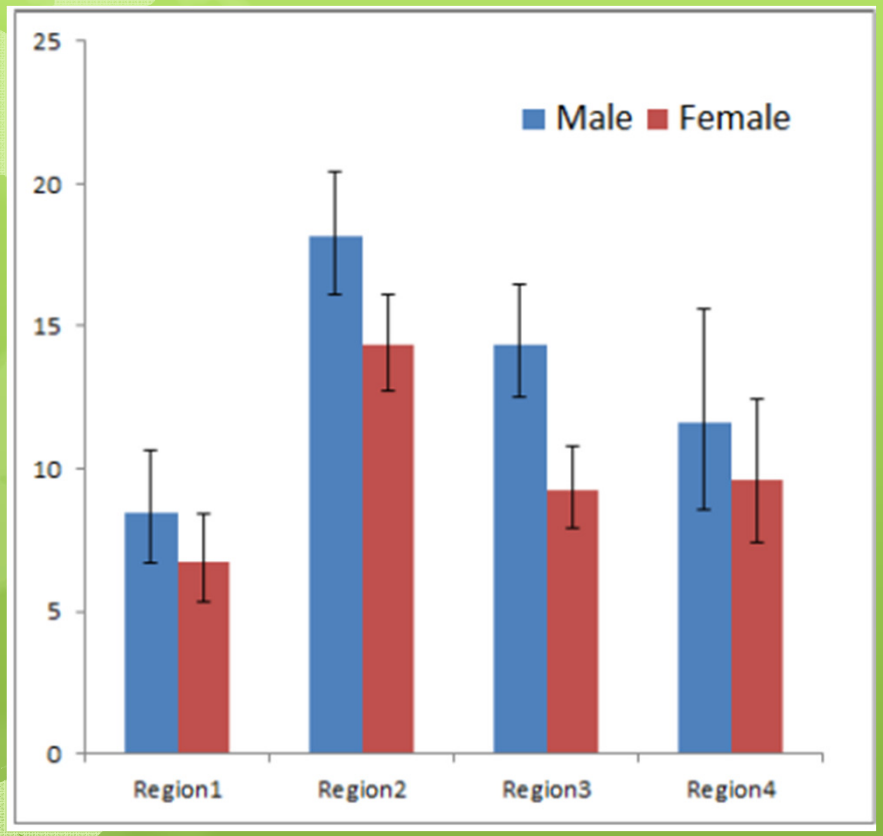
- Look for overlap of error lines:
  - If they **overlap**: The difference is **NOT** significant
  - If they **don't overlap**: The difference **may be** significant





Which is a valid statement?

- Fish2Whale food caused the most fish growth
- Fish2Whale food caused more fish growth than did Budget Fude



Statements:

- In all four regions, more males exhibited the trait measured than did females.

- More males in region 3 exhibited the measured trait than did females